Appl. No. 10/055,352
Appeal Brief in Response
to final Office action of 8 October 2004

RECEIVED
CENTRAL FAX CENTER
MAR 0 3 2005

IN THE UNITED STATES

PATENT AND TRADEMARK OFFICE

Appl. No.	: **10/055,352**
Applicant(s)	: **GUTTA et al.**
Filed	: **1/22/2002**
TC/A.U.	: **2121**
Examiner	: **HIRL, Joseph P.**
Atty. Docket	: **US020030**

CERTIFICATE OF MAILING OR TRANSMISSION

I certify that this correspondence is being:
[ ] deposited with the U.S. Postal Service with sufficient postage as first-class mail in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.
[X] transmitted by facsimile to the U.S. Patent and Trademark Office at **703-872-9306**

On: 3 March 2005

By: [signature]

Title: **SELF-ORGANIZING FEATURE MAP WITH IMPROVED PERFORMANCE BY NON-MONOTONIC VARIATION OF THE LEARNING RATE**

Mail Stop: **APPEAL BRIEF - PATENTS**
Commissioner for Patents
Alexandria, VA 22313-1450

APPEAL UNDER 37 CFR 41.37

Sir:

This is an appeal from the decision of the Examiner dated 8 October 2004, finally rejecting claims 1-14 of the subject application.

This paper includes (each beginning on a separate sheet):

1. **Appeal Brief;**
2. **Claims on Appeal; and**
3. **Credit card authorization in the amount of $500.**

03/04/2005 EKOLI1 00000002 10055352

01 FC:1402 500.00 OP

APPEAL BRIEF

I. REAL PARTY IN INTEREST

The above-identified application is assigned, in its entirety, to **Koninklijke Philips Electronics N. V.**

II. RELATED APPEALS AND INTERFERENCES

Appellant is not aware of any co-pending appeal or interference which will directly affect or be directly affected by or have any bearing on the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1-14 are pending in the application.

Claims 1-14 stand rejected by the Examiner under 35 U.S.C. 101.

Claims 1-14 stand rejected by the Examiner under 35 U.S.C. 102(b).

These rejected claims are the subject of this appeal.

IV. STATUS OF AMENDMENTS

No amendments were filed subsequent to the final rejection in the Office Action dated 8 October 2004.

V. SUMMARY OF CLAIMED SUBJECT MATTER

As claimed in independent claim 1, the invention comprises a method for training a self ordering map for use in a computing system. The method includes initializing a set of weights of the self ordering map and iteratively training the weights over many training epochs. For at least a number of the training epochs, the iterative training includes updating the weights based on a learning rate that is generated according to a function that changes in a fashion that is other than monotonically decreasing with the training epochs.

As discovered by the applicants, by changing the learning rate in other than a monotonically decreasing fashion, the rate of convergence to a solution often increases, compared to the conventional method of using a monotonically decreasing function (Applicants' page 7, line 18 through page 8, line 5).

The applicants' FIG. 2 is an example flow diagram of this invention. The weights are initialized, at S10, preferably with randomly generated values. A sample input vector is drawn from a pool of training vectors, at S20, and a winning node is identified, at S30. An example method for determining a winning node is presented at page 9, line 21 through page 10, line 7 of the applicants' specification.

The self ordering map is trained based on the distance of each node from the winning node, so that nodes far from the winning node are updated more strongly than nodes closer to the winning node. The degree of updating of all the nodes is termed the learning rate; a high learning rate causes larger magnitude changes than a lower learning rate. In the example formula at page 10, line 12, the variable α is the learning rate, and it is used to linearly scale the amount of change (ΔW) of each node's weight.

At S40 of FIG. 2, the learning rate is determined for the current training cycle, and at S50, the weights are accordingly updated. In accordance with this invention, the learning rate does not monotonically decrease. That is, at some point in the learning process, the learning rate is greater than a prior learning rate. As noted by the applicants, by allowing subsequent training epochs to have a greater effect on the learning than prior epochs, experiments have shown that convergence is often achieved more quickly (page 7, line 18 through page 8, line 5).

Convergence is generally determined when subsequent training epochs fail to produce a substantial change in the weights, as illustrated by the test at S60 of FIG. 2, and presented at page 10, lines 16-18 of the applicants' specification.

As illustrated in the applicants' example FIG. 3, in a preferred embodiment, the learning rate generally decreases with each training epoch, but, in accordance with this invention, is permitted to vary between upper and lower bounds 161-162. As can be seen, by allowing the learning rate to vary between these bounds 161-162, a subsequent epoch can have a higher learning rate than a prior epoch. For example, at one epoch, the learning rate

may lie on curve 162, and at the next epoch, the learning rate may lie on curve 161, which is above the curve 162.

The use of higher-than-prior learning rates has been determined to be particularly effective in the initial training epochs, and less effective as the training approaches convergence. In the example presented at FIG. 4, the variance of the learning rate is limited to the initial training epochs (page 11, lines 19-23).

A monotonically decreasing function never increases. A subsequent epoch in a training system with a monotonically decreasing learning rate can never exhibit a higher learning rate than a prior epoch. Thus, a characterization of the applicants' sometimes-larger change to the learning rate is properly defined as "other than monotonically decreasing".

As claimed in independent claim 8, the invention comprises a method of training a self ordering feature map for use in a computing system, comprising: choosing a random value for initial weight vectors; drawing a sample from a set of training sample vectors and applying it to input nodes of the self ordering feature map; identifying a winning competition node of the self ordering feature map according to a least distance criterion; adjusting a synaptic weight of at least the winning node, using a learning rate to update the synaptic weight that is based on a function other than one that is monotonic with subsequent training epochs; iteratively repeating the drawing, identifying, and adjusting to form each subsequent training epoch.

Claim 8 includes the elements of the example method illustrated in FIG. 2, discussed above. Of particular note, claim 8 includes using a learning rate to update the synaptic weight that is based on a function other than one that is monotonic with subsequent training epochs, as presented at page 7, line 18 through page 8, line 5, and at page 11, lines 3-23.

VI. ISSUES TO BE REVIEWED ON APPEAL

Claims 1-14 stand rejected under 35 U.S.C. 101.

Claims 1-14 stand rejected under 35 U.S.C. 102(b) over Mehrotra (MIT Press, 1997, Artificial Neural Networks).

VII. ARGUMENT

Rejection under 35 U.S.C. 101

35 U.S.C. 101 states:

"Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title."

Claims 1-14

The Office action asserts that "a sporadic alteration of the learning rate does not *always* increase the rate of convergence. Consequently, for the lack of concreteness, the disclosure is non statutory". The applicants note that the Office action has not challenged the applicants' statement that a sporadic alteration of the learning rate *often* increases the rate of convergence, but rather, the Office action demands that the sporadic alteration must *always* increase the rate of convergence.

The applicants respectfully note that absolute success and/or efficiency, as demanded by the Office action, is not a criterion for patentability under 35 U.S.C. 101. The applicants further note that many useful computer processes are not optimal under all circumstances, just as many useful mechanical devices are not efficient under all circumstances, and many useful manufacturing processes do not result in cost and/or time savings under all circumstances. Most classical mathematical/computer problems, even ones as simple as sorting and routing, do not yet have a "perfect" solution that is guaranteed to provide a better/more-efficient result than all other solutions, even after decades, and in some cases centuries, of research. In the relatively newer field of machine learning, to which this invention is addressed, the situation is no better. To deny a patent to all non-perfect inventions is contrary to the basic spirit and intent of U.S. patent laws.

MPEP 2106 specifically provides guidance for evaluating computer-related inventions:

> "Office personnel have the burden to establish a *prima facie* case that the claimed invention as a whole is directed to solely an abstract idea or to manipulation of abstract ideas or does not produce a useful result. ***Only when the claim is devoid of any limitation to a practical application in the technological arts should it be***

rejected under 35 U.S.C. 101. Compare *Musgrave*, 431 F.2d at 893, 167 USPQ at 289; *In re Foster*, 438 F.2d 1011, 1013, 169 USPQ 99, 101 (CCPA 1971). Further, when such a rejection is made, Office personnel must expressly state how the language of the claims has been interpreted to support the rejection. ...

As the Supreme Court has held, Congress chose the expansive language of 35 U.S.C. 101 so as to include "***anything under the sun that is made by man.***" *Diamond v. Chakrabarty*, 447 U.S. 303, 308-09, 206 USPQ 193, 197 (1980). Accordingly, section 101 of title 35, United States Code, provides:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

In *Chakrabarty*, 447 U.S. at 308-309, 206 USPQ at 197, the court stated:

In choosing such expansive terms as "manufacture" and "composition of matter," modified by the comprehensive "any," Congress plainly contemplated that the patent laws would be given wide scope. The relevant legislative history also supports a broad construction. The Patent Act of 1793, authored by Thomas Jefferson, defined statutory subject matter as "any new and useful art, machine, manufacture, or composition of matter, or any new or useful improvement [thereof]." Act of Feb. 21, 1793, ch. 11, § 1, 1 Stat. 318. The Act embodied Jefferson's philosophy that "ingenuity should receive a liberal encouragement." V Writings of Thomas Jefferson, at 75-76. See *Graham* v. *John Deere Co.*, 383 U.S. 1, 7-10 (148 USPQ 459, 462-464) (1966). Subsequent patent statutes in 1836, 1870, and 1874 employed this same broad language. In 1952, when the patent laws were recodified, Congress replaced the word "art" with "process," but otherwise left Jefferson's language intact. The Committee Reports accompanying the 1952 Act inform us that Congress intended statutory subject matter to "include anything under the sun that is made by man." S. Rep. No. 1979, 82d Cong., 2d Sess., 5 (1952); H.R. Rep. No. 1923, 82d Cong., 2d Sess., 6 (1952). [Footnote omitted]

This perspective has been embraced by the Federal Circuit:

The plain and unambiguous meaning of section 101 is that ***any new and useful process***, machine, manufacture, or composition of matter, or any new and useful improvement thereof, ***may be patented*** if it meets the requirements for patentability set forth in Title 35, such as those found in sections 102, 103, and 112. The use of the expansive term "any" in section 101 represents Congress's intent not to place any restrictions on the subject matter for which a patent may be obtained beyond those specifically recited in section 101 and the other parts of Title 35. . . . Thus, **it is improper to read into section 101 limitations as to the subject matter that may be patented where the legislative history does not indicate that Congress clearly intended such limitations.** *Alappat*, 33 F.3d at 1542, 31 USPQ2d at 1556.

The applicants respectfully note that Congress did not say "any thing under the sun that is made by man *that guarantees improvement under all circumstances*", as the Office action implies. The applicants also note that MPEP 2106 states "Only when the claim is devoid of any limitation to a practical application in the technological arts should it be

rejected", and does not state "Whenever the claim does not guarantee an improvement under all circumstances it should be rejected", as the Office action implies.

Because each of the rejected claims includes a new and useful process that has practical application in the technical arts, the applicants respectfully maintain that the rejection of claims 1-14 under 35 U.S.C. 101 is unfounded, and not in accordance with the specific directives of MPEP 2106.

Rejection under 35 U.S.C. 102(b) over Mehrotra

MPEP 2131 states:

> "A claim is anticipated only if ***each and every element*** as set forth in the claim is found, either expressly or inherently described, in a single prior art reference." *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). "The ***identical invention*** must be shown in as ***complete detail*** as is contained in the ... claim." *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989).

Claims 1-7

Claim 1, upon which claims 2-7 depend, claims a method for training a self ordering map that includes updating weights of the map based on a learning rate that is generated according to a function that changes in a fashion that is other than monotonically decreasing with the training epochs. Mehrotra fails to teach updating weights based on a learning rate that is generated according to a function other than monotonically decreasing, as specifically claimed.

Mehrotra specifically teaches a monotonically decreasing function for updating the learning rate at each epoch, at page 192, lines 1-3, wherein a step function is used to reduce the learning rate. The Office action acknowledges that Mehrotra's function does not increase, but maintains that because Mehrotra's function includes level portions, it is not a monotonically decreasing function. The applicants respectfully disagree with this interpretation of "monotonically decreasing".

The Office action erroneously defines monotonically decreasing to mean "never remaining constant or increasing". The applicants respectfully note that this is the definition of "strictly decreasing", and not "monotonically decreasing". The Office action's definition renders the term "monotonic" superfluous, because "never remaining constant or increasing" is the literal/strict definition of "decreasing".

The applicants have cited Webster's definition of monotonic as "having the property of never increasing or never decreasing as the independent variable increases". The Office action cites "Mathworld.wolfram.com"[1] for teaching: "A function which is either entirely nonincreasing or nondecreasing." The applicants respectfully note that these references use negative terms ("never increasing", "never decreasing", "entirely nonincreasing", and "entirely nondecreasing") to define monotonicity, as contrast with the simpler, but different, affirmative terms "always decreasing", "always increasing", "entirely increasing", and "entirely decreasing" that would conform to the definition asserted in the Office action.

The applicants respectfully maintain that the reason these references use the more cumbersome negative terms ("nonincreasing" in lieu of "decreasing"; "nondecreasing" in lieu of "increasing") is that monotonic functions may include a level state. That is, a "nonincreasing" function includes decreasing values as well as level values, and a "nondecreasing" function includes both increasing values and level values.

The applicants further note that in the reference cited by the Office action, "Mathworld.wolfram.com", the terms "nonincreasing" and "nondecreasing" in the definition of monotonic are hypertext items, and lead to the following definition:

> A function $f(x)$ is said to be nonincreasing on an interval I if $f(b) \leq f(a)$ for all $b > a$, where $a, b \in I$. Conversely, a function $f(x)$ is said to be nondecreasing on an interval I if $f(b) \geq f(a)$ for all $b > a$ with $a, b \in I$.

The applicants respectfully note the use of the "less than ***or equal to***" sign ($\leq$) and the "greater than ***or equal to***" sign ($\geq$) in the above formal mathematical definition of "nonincreasing" and "nondecreasing" as used in the definition of a monotonic function. That is, regions where $f(b) = f(a)$ for $b>a$ (the level steps of Mehrotra) are included in the mathematical definition of a monotonic function, thereby verifying that Mehrotra teaches a monotonic function.

[1] The applicants' remarks herein should not be interpreted as an endorsement of the use of a web-page as a technical reference. Rarely do web pages undergo the rigor of formal reference text. "Not everything that's published on the Web is true."

The Office action also cites "Mathworld.wolfram.com" for further defining a "function is monotonic if its first derivative does not change sign". The applicants respectfully maintain that the value zero is unsigned, and thus a function that includes level values (derivative of zero) cannot be said to "change sign" when its derivative goes to zero.

Although the applicants do not necessarily endorse the use of web-page references, the following cites are provided in an attempt to further clarify the term "monotonic".

Dictionary.com defines:

> **mon·o·ton·ic** (mŏn'ə-tŏn'ĭk) *Mathematics.* Designating sequences, the successive members of which either consistently increase or decrease but do not oscillate in relative value. Each member of a monotone increasing sequence is greater than ***or equal to*** the preceding member; each member of a monotone decreasing sequence is less than ***or equal to*** the preceding member.

Wikipedia.org provides a clear distinction between "montonically decreasing/increasing" and "strictly decreasing/increasing":

> If the order ≤ in the definition of ***monotonicity*** is replaced by the strict order <, then one obtains a stronger requirement. A function with this property is called **strictly increasing**. Again, by inverting the order symbol, one finds a corresponding concept called **strictly decreasing**.

Because Mehrotra specifically teaches a monotonically decreasing function for updating the learning rate at each epoch, and the applicants specifically claim a function that is other than monotonically decreasing, the applicants respectfully maintain that claims 1-7 are patentable under 35 U.S.C. 102(b) over Mehrotra.

Claims 8-14

Claim 8, upon which claims 9-14 depend, claims a method of training a self ordering feature map that includes using a learning rate to update the synaptic weight that is based on a function other than one that is monotonic with subsequent training epochs. Mehrotra fails to teach using a learning rate that is based on a function that is other than monotonic.

As noted above, Mehrotra specifically teaches a monotonic function for reducing the learning rate that is used to update the synaptic weights of a self ordering feature map.

Because Mehrotra fails to teach a function that is other than monotonic for updating the synaptic weight, as specifically claimed by the applicants, the applicants respectfully maintain that claims 8-14 are patentable under 35 U.S.C. 102(b) over Mehrotra.

CONCLUSIONS

Because claims 1-14 recite a new and useful process that is applicable to the technical arts, the applicants respectfully request that the Examiner's rejection of claims 1-14 under 35 U.S.C. 101 be reversed by the Board, and the claims be allowed to pass to issue.

Because claims 1-14 specifically claim changing a learning rate in other than a monotonic fashion, and Mehrotra specifically teaches a monotonic change to the learning rate, the applicants respectfully request that the Examiner's rejection of claims 102(b) under 35 U.S.C. 102(b) be reversed by the Board, and the claims be allowed to pass to issue.

Respectfully submitted,

Robert M. McDermott, Attorney
Registration Number 41,508
804-493-0707

APPENDIX

CLAIMS ON APPEAL

1. A method for training a self ordering map for use in a computing system, comprising:

initializing a set of weights of the self ordering map; and

iteratively training the weights over many training epochs;

wherein

for at least a number of the training epochs, iteratively training the weights includes

updating the weights based on a learning rate that is generated according to a function that changes in a fashion that is other than monotonically decreasing with the training epochs.

2. A method as in claim 1, wherein

the function includes a random or pseudorandom function.

3. A method as in claim 2 wherein

the random or pseudorandom function has a range that decreases with the training epochs.

4. A method as in claim 2 wherein

the random or pseudorandom function is configured such that the learning rate tends to decrease with the training epochs.

5. A method as in claim 1 wherein

the function has a range that decreases with the training epochs.

6. A method as in claim 5 wherein

the function is configured such that the learning rate tends to decrease with the training epochs.

7. A method as in claim 1 wherein

the function is configured such that the learning rate tends to decrease with the training epochs.

8. A method of training a self ordering feature map for use in a computing system, comprising

choosing a random value for initial weight vectors;

drawing a sample from a set of training sample vectors and applying it to input nodes of the self ordering feature map;

identifying a winning competition node of the self ordering feature map according to a least distance criterion;

adjusting a synaptic weight of at least the winning node, using a learning rate to update the synaptic weight that is based on a function other than one that is monotonic with subsequent training epochs;

iteratively repeating the drawing, identifying, and adjusting to form each subsequent training epoch.

9. A method as in claim 8, wherein

the function corresponds to a random or pseudorandom function.

10. A method as in claim 9 wherein

the function has a range that decreases with subsequent training epochs.

11. A method as in claim 9 wherein

the function is configured such that the learning rate tends to decrease with subsequent training epochs.

12. A method as in claim 8 wherein

the function has a range that decreases with subsequent training epochs.

13. A method as in claim 12 wherein

the function is configured such that the learning rate tends to decrease with subsequent training epochs.

14. A method as in claim 8 wherein

the function is configured such that the learning rate tends to decrease with subsequent training epochs.